

# SCIENTIFIC RESEARCH LABORATORIES

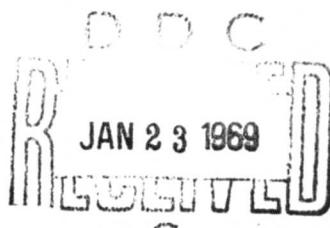
AD 680811

## Some Practical Regularity Conditions for Nonlinear Programs

David W. Walkup

Roger J.-B. Wets

This document has been approved  
for public release and sale; its  
distribution is unlimited.



Photocopies by the  
CLEARINGHOUSE  
for Federal Scientific & Technical  
Information, Springfield, Va. 22151

MATHEMATICS RESEARCH

NOVEMBER 1968

D1-82-0792

SOME PRACTICAL REGULARITY CONDITIONS FOR NONLINEAR PROGRAMS

by

David W. Walkup

and

Roger J.-B. Wets

Mathematical Note No. 583

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

November 1968

#### ABSTRACT

Some practical sufficient conditions are given for a program with linear constraints and nonlinear objective to have well-behaved duality properties. Thus if the objective is convex, lower semicontinuous, and satisfies a Lipschitz condition on a closed convex polyhedron, and is  $+\infty$  elsewhere, then there exists a dual problem which actually achieves the optimal value of the primal problem. An argument is given suggesting that certain of the results are best possible from the viewpoint of practical applications. These results are applied to show that a broad class of stochastic programs with recourse have desirable duality properties.

## 1. INTRODUCTION

Theorem 1 below gives sufficient conditions for a program with linear constraints and a convex objective with range in the extended reals to have well-behaved duality properties. The principal ingredient in the proof of Theorem 1 is Theorem 2, which considers the properties of the optimum value of a (not necessarily convex) program under perturbation of the linear constraints. The terminology used in Theorem 1 is adapted from [1,2,3] where the close relationship between properties of the variational function and the duality properties of a program have been discussed at some length. Briefly a program is *solvab*le if the value of the infimum is finite and achieved for some value of the variables, it is *dualizable* if there is no duality gap, and it is *stable* if there exist (optimal) nontrivial Lagrange multipliers. The definitions of convexity and lower semicontinuity for functions into the extended reals will be reviewed in the next section.

At the end of §2 an argument will be given suggesting that Theorem 1 is the best possible from the viewpoint of practical applications. Finally, Theorem 1 will be applied in §4 to show that a broad class of stochastic programs with recourse have desirable duality properties.

THEOREM 1. Consider the nonlinear program

$$\begin{aligned} \inf \quad & f(x) \\ \text{Ax} = b \\ x \geq 0 \end{aligned} \tag{1}$$

where the objective  $f$  is convex in the sense of functions into the extended reals.

- (i) If  $f$  is lower semicontinuous in the sense of functions into the extended reals and the constraint set  $K = \{x | Ax = b, x \geq 0\}$  is bounded and contains at least one point where  $f$  is finite, then (1) is solvable and dualizable.
- (ii) If  $f$  is finite and Lipschitz on a closed convex polyhedron and (1) has a finite value, then (1) is stable.

THEOREM 2. Consider the function  $\phi(u) = \inf\{f(x) | x \in \kappa(u)\}$ , where  $\kappa(u) = \{x | Ax = b - u, x \geq 0\}$  and  $f$  is a function with range in the extended reals.

- (i) If  $f$  is convex, so is  $\phi$ .
- (ii) If  $f$  is lower semicontinuous and  $\kappa(u)$  is compact for some  $u$ , then  $\phi$  is lower semicontinuous.
- (iii) If  $f$  is  $+\infty$  except on some closed convex polyhedron where it is either finite and Lipschitz or identically  $-\infty$ , then the same holds for  $\phi$ .

## 2. PROOF OF THEOREM 2

DEFINITIONS. Let  $f$  be a function with  $R^n$  for domain and the extended reals  $\bar{R} = R \cup \{+\infty\} \cup \{-\infty\}$  for range. The set

$$\text{epi } f = \{(z, x) | z \in R, x \in R^n, z \geq f(x)\}$$

is called the epigraph of  $f$ . The function  $f$  is said to be convex if its epigraph is a convex subset of  $R^{n+1}$  or equivalently if

$$f(x_\lambda) = f[(1-\lambda)x_0 + \lambda x_1] \leq (1-\lambda)f(x_0) + \lambda f(x_1)$$

for all  $\lambda \in [0,1]$  and  $x_0, x_1 \in \mathbb{R}^n$ , where the conventions  $0 \cdot \infty = 0$  and  $(+\infty) + (-\infty) = +\infty$  apply. The function  $f$  is said to be lower semicontinuous if

$$\liminf f(x_i) \geq f(\lim x_i)$$

for every convergent sequence  $(x_i)$  in  $\mathbb{R}^n$ , or equivalently if  $\text{epi } f$  is a closed subset of  $\mathbb{R}^{n+1}$ . If the function  $f$  is finite on some subset  $S$  of  $\mathbb{R}^n$  and  $B$  is a real number such that

$$|f(x) - f(x')| \leq B ||x - x'||$$

for all  $x, x'$  in  $S$ , where  $||\cdot||$  is the Euclidean norm in  $\mathbb{R}^n$ , then  $f$  is said to be Lipschitz on  $S$  with constant  $B$ .

The following two lemmas represent an application of results in [6] to the special situation of Theorem 2. (An inconsequential difference is that in [6]  $\kappa(u)$  is a section of  $P$  rather than its projection into  $\mathbb{R}^n$ .)

LEMMA 1. Let  $P = \{(x, u) | Ax + Du = b, x \geq 0\}$ ,  $\kappa(u) = \{x | Ax = b - Du, x \geq 0\}$ ,  $Q = \{u | \kappa(u) \text{ is nonempty}\}$ , and let  $P = P_0 + C$  be the representation of the polyhedron  $P$  as the vector sum of a bounded polyhedron  $P_0$  and a (unique) polyhedral cone  $C$  with apex at the origin. Then  $Q$  is a polyhedron, and for each  $u \in Q$ ,  $\kappa(u) = \kappa_0(u) + C'$ , where  $\kappa_0(u)$  is a bounded polyhedron depending on  $u$  and  $C'$  is the polyhedral cone  $\{x | (x, 0) \in C\}$ . In particular,  $\kappa(u)$  is bounded for all  $u$  in  $Q$  if it is bounded for some  $u$  in  $Q$ , for in this case  $C' = \{0\}$ .

LEMMA 2. Let  $\kappa$  and  $Q$  be defined as in Lemma 1. Then there exists a constant  $\bar{B}$  such that for any two points  $u_1$  and  $u_2$  in  $Q$

$$d[\kappa(u_1), \kappa(u_2)] \leq \bar{B} \|u_1 - u_2\|,$$

where  $d[ , ]$  denotes the Hausdorff distance between sets in  $R^n$ .

Observe that Lemmas 1 and 2 apply directly to Theorem 2 if  $u$  is taken to be a vector in  $R^m$  and the matrix  $D$  is the identity. Accordingly let  $P$  and  $Q$  be defined in this manner, and consider part (i) of Theorem 2. Now  $P$  is convex, and by assumption  $\text{epi } f$  is a convex subset of  $R \times R^n$ . Hence  $(\text{epi } f \times R^m) \cap R \times P$  is a convex subset of  $R \times R^n \times R^m$  and its projection onto the space  $R \times R^m$  is a convex set  $C$ . But  $\text{epi } \phi$  is just the vertical closure of  $C$ , i.e., the union of  $C$  and any missing endpoints of "vertical" line segments in  $C$ . This proves part (i) of Theorem 2. Also,  $Q$  is closed and  $\phi(u) = +\infty$  if  $u$  is not in  $Q$ . Thus, in order to prove part (ii) of Theorem 2 it suffices to show

$$\liminf \phi(u_i) \geq \phi(u_0),$$

where  $u_i \in Q$  and  $\lim u_i = u_0 \in Q$ . By Lemma 1 each  $\kappa(u)$  is compact, and hence the lower semicontinuous function  $f$  attains a minimum on  $\kappa(u)$  at some point, say  $x(u)$ . Also, it follows from Lemma 2 that  $\kappa(u)$  is uniformly bounded for  $u$  in some compact neighborhood  $N$  of  $u_0$ , i.e.,  $\kappa \cap [R^n \times N]$  is compact. Hence the sequence of points  $x(u_i)$  has at least one limit point  $x_0 \in \kappa(u_0)$ . Then

$$\liminf \phi(u_i) = \liminf f(x(u_i)) \geq f(x_0) \geq f(x(u_0)) = \phi(u_0),$$

where the first inequality follows from the lower semicontinuity of  $f$  and the second follows from the optimality of  $x(u_0)$  on  $\kappa(u_0)$ .

This proves part (ii) of Theorem 2.

We shall begin the proof of part (iii) of Theorem 2 by establishing the following special case:

LEMMA 3. Under the conditions of Theorem 2, if  $f$  is finite and Lipschitz throughout  $R^m$ , then either  $\phi$  is identically  $-\infty$  on  $Q$  or  $\phi$  is finite and Lipschitz on  $Q$ .

Proof. Let  $u$  and  $u'$  be any two points of  $Q$ , let  $(x_i)$  be a sequence of points in  $\kappa(u)$  such that  $\lim i(x_i) = \phi(u)$ , and let  $x'_i$  be the point of  $\kappa(u')$  closest to  $x_i$ . Then

$$\begin{aligned} \phi(u') - f(x_i) &\leq f(x'_i) - f(x_i) \\ &\leq B ||x'_i - x_i|| \\ &\leq \bar{B} ||u' - u|| \end{aligned} \tag{2}$$

where  $B$  is the Lipschitz constant for  $f$  and  $\bar{B}$  is the constant of Lemma 2. Now  $\phi(u')$  may be  $-\infty$  or finite, and  $\lim f(x_i) = \phi(u)$  may be  $-\infty$  or finite. But (2) shows that  $\phi(u) = -\infty$  implies  $\phi(u') = -\infty$ , i.e.,  $\phi$  is identically  $-\infty$  on  $Q$  or  $\phi$  is finite on  $Q$ . And if  $\phi$  is finite on  $Q$ , then (2) implies  $\phi(u') - \phi(u) \leq \bar{B} ||u' - u||$ . By the symmetry in  $u$  and  $u'$  it follows that  $\phi$  is Lipschitz with constant  $\bar{B}$ .

Now suppose, as assumed in Theorem 2, that the range of  $f$  is

contained in the extended reals, and the set  $K = \{x | f(x) < +\infty\}$  is a closed convex polyhedron. Then  $\phi$  may be defined by the program

$$\begin{aligned}\phi(u) &= \inf f(x) \\ Ax &= b - u \\ A'x &\geq 0 \\ x &\in K\end{aligned}\tag{3}$$

Since  $K$  is a polyhedron, (3) may be rewritten

$$\begin{aligned}\phi(u) &= \inf f(x) + 0x' \\ Ax &= b - u \\ A'x + x' &= b' \\ x &\geq 0, x' \geq 0\end{aligned}\tag{4}$$

The set  $K_0 = \{u | \phi(u) < +\infty\}$  is exactly the set of  $u$  for which the constraints of (3) or (4) are feasible, hence this set is a polyhedron. Moreover, if  $f$  is  $-\infty$  on  $K$ , then  $\phi$  is  $-\infty$  on  $K_0$ . This establishes a portion of part (iii) of Theorem 2. The remaining possibility is that  $f$  is finite and Lipschitz on  $K$ . Now  $f$  may not be finite and Lipschitz where (4) is infeasible, and  $u$  does not perturb all constraints of (4), but clearly the proof of Lemma 3 will apply to (4), with a different choice of  $D$  in Lemma 1, and yield the remainder of Theorem 2.

Strictly speaking the requirement that  $\{x | f(x) < +\infty\}$  be a polyhedron is not essential to Theorem 2. The proof of Lemma 3 clearly establishes the following more general result.

LEMMA 4. Suppose  $f$  is finite and Lipschitz everywhere,  $K$  is a closed convex subset of  $R^n \times R^m$ ,  $\phi(u) = \inf \{f(x) | x \in \kappa(u)\}$ ,  $\kappa(u) = \{x | (x, u) \in K\}$ , and  $Q = \{u | \kappa(u) \text{ is nonempty}\}$ . Then  $\phi$  is either  $-\infty$  on  $Q$  or finite and Lipschitz on  $Q$  provided there exists a constant  $\bar{B}$  such that

$$d[\kappa(u), \kappa(u')] \leq \bar{B} ||u - u'|| \quad (5)$$

for all  $u, u'$  in  $Q$ .

However, it is difficult to see how this more general result can be applied to practical problems. For consider a program of the form

$$\begin{aligned} \phi(u) &= \inf f(x) \\ Ax &= b - u \\ x &\in K \end{aligned}$$

where  $K$  is a convex set. Surely any practical condition to be imposed on  $P$  ought not to depend on the particular form of  $A$ . But it is shown in [6] that if  $P$  is not a polyhedron there even exist  $a_1, \dots, a_n$  such that

$$\kappa(u_1) = \{x | a_1 x_1 + \dots + a_n x_n = -u_1, x \in P\}$$

fails to satisfy (5) for any  $\bar{B}$ .

### 3. DERIVATION OF THEOREM 1

For easy reference, we repeat the convex program of Theorem 1:

$$\begin{aligned} & \text{Inf} \quad f(x) \\ & Ax = b \\ & x \geq 0 \end{aligned} \tag{1}$$

The basic duality properties of this program are conveniently represented in an equivalent *Inf* problem

$$\begin{aligned} & \text{Inf} \quad n \\ & (n, u) \in \mathcal{L} \cap \mathcal{C} \end{aligned} \tag{6}$$

where  $\mathcal{L}$  is the vertical line  $\{(n, u) | n \in \mathbb{R}, u = 0\}$  and  $\mathcal{C}$  is the convex set whose vertical closure is the epigraph of  $\phi$ , as described in the proof of part (i) of Theorem 1. The program (1) will be *solvable*, i.e., have a finite optimum value  $\phi(0) = f(x^0)$  achieved by some feasible  $x^0$ , if and only if  $\mathcal{L} \cap \mathcal{C}$  is a closed half line.

A natural dual to the *Inf* problem (6) is the *Sup* problem:

$$\begin{aligned} & \text{Sup} \quad \mu \\ & (\mu, u^*) \\ & \mu \leq n + u^*u \text{ for all } (n, u) \in \mathcal{C}. \end{aligned}$$

The *Inf* problem asks for the infimal height  $n$  of points on  $\mathcal{L}$  inside  $\mathcal{C}$ . The *Sup* problem asks for the supremal height  $\mu$  of points on  $\mathcal{L}$  through which can be passed nonvertical hyperplanes (i.e., those not containing any line parallel to  $\mathcal{L}$ ) bounding  $\mathcal{C}$ . The *Sup* problem can be recast in the form

$$\begin{aligned} & \text{Sup} \quad \mu \\ & (\mu, u^*) \\ & \mu \leq f(x) + u^*(b - Ax) \text{ for all } x \geq 0 \end{aligned}$$

where the role of  $u^*$  as a row  $m$ -vector of Lagrange multipliers is

apparent.

It is natural to say that the Inf and Sup problems are *dual* or that (1) is *dualizable* if  $\text{Sup } \mu = \text{Inf } \eta$ , i.e., if there is no duality gap. We shall say that the program (1) is *stable* if the Sup problem is solvable, i.e., has a finite optimum which is achieved for some  $u^*$ . Thus (1) is stable if there is a nonvertical hyperplane supporting  $C$  at the point  $(\phi(0), 0)$  on  $L$ . Equivalently, (1) is stable if there exist Lagrange multipliers which convert it into an equivalent unconstrained problem.

The above discussion of duality properties of (1) follows the approach to mathematical programming in abstract spaces given in [3] which is closely related to the more extensive development given by Rockafellar in [1,2]. Theorem 1 now follows easily from the above discussion, well-known properties of convex sets and their supports, the simple relationship between  $C$  and the epigraph of  $\phi$ , and from Theorem 2.

REMARK. We have seen that (1) is stable if  $f$  satisfies a Lipschitz condition. But also, the duality theory for quadratic programs shows that (1) is stable if  $f$  is convex and quadratic, and hence definitely not Lipschitz. We do not know of any sensible condition on  $f$  which implies both properties and is valid in Theorem 1. A simple example will show that (1) may fail to be stable if  $K$  is unbounded and  $f$  is only locally Lipschitz on a polyhedron.

#### 4. APPLICATIONS TO STOCHASTIC PROGRAMS WITH RE COURSE

In this section we will use Theorem 1 to show that fairly broad classes of stochastic programs with recourse have well-behaved duals. The necessary results on the objective functions of such programs have already been proved in earlier papers [4,5]. We shall make liberal use of the notation and terminology introduced in [4,5,8].

THEOREM 3. If the value of a stochastic program with recourse is finite and the first-stage feasibility set  $K_1 = \{x | Ax = b, x \geq 0\}$  is bounded, then the equivalent convex program is solvable. Moreover, either the equivalent convex program is dualizable or the objective takes the value  $-\infty$  at points belonging to  $K_1$  for arbitrarily small perturbation of  $b$ , i.e., there is an infinite duality gap.

Proof. It is shown in [5] that the objective  $z(x)$  of the equivalent convex program is either lower semicontinuous as a function into the extended reals or takes the value  $-\infty$  at some point. The second part of the theorem follows immediately from this and Theorem 1. Now let  $K_2^S$  be the set on which  $z(x)$  is less than  $+\infty$  and let  $M$  be the affine hull of the intersection of  $K_1$  and  $K_2^S$ . Since  $z$  is convex, so is  $K_2^S$ , and hence  $K_1 \cap K_2^S$  has an interior with respect to  $M$ . Since the restriction  $z_M$  of  $z$  to  $M$  is convex, and since it is finite on  $K_1 \cap K_2^S$ , it follows that it is nowhere  $-\infty$ . The results of [5] apply equally well to  $z_M$  and show that it is lower semicontinuous. A straightforward application of Theorem 1 completes the proof.

THEOREM 4. If the value of a stochastic program with recourse is finite, the recourse matrix  $W$  is fixed, the second-stage feasibility set  $K_2$  is a polyhedron, and the random variables  $\xi$  have finite variance, then the equivalent convex program is stable.

Proof. Theorem (4.5) of [4] shows that under the second and fourth hypotheses either the objective of the equivalent convex program is  $-\infty$  throughout  $K_2$  or it is finite and Lipschitz on  $K_2$ . The rest follows from Theorem 1.

It is worth mentioning that Proposition (3.16) of [4] gives practical conditions which insure that  $K_2$  is a polyhedron. In addition Corollary (4.7) of [4] gives some alternate conditions on the distribution of  $\xi$  under which the conclusions of Theorem (4.5), and hence the above theorem, remain valid.

We conclude with a few remarks showing that the hypotheses of Theorems 3 and 4 are not superfluous. Consider the special class of stochastic programs considered in detail in [7], for which  $W = [I, -I]$  and only the right-hand sides are random. It is easy to construct examples in this class such that the random variables have a distribution possessing all moments, and the value of the infimum is finite, but the program is not solvable, provided  $K_1$  is allowed to be unbounded. Thus boundedness of  $K_1$  is not superfluous for solvability in Theorem 3. We remark that in such an example  $z(x)$  will be finite everywhere, and hence by Theorem 4 such an example will be stable, but not solvable.

A stochastic program can be constructed with  $T$  only random and  $K_1$  bounded for which the alternative at the end of Theorem 3 occurs. An example can also be constructed with  $T$  only random and  $K_1$  unbounded for which  $z(x) > -\infty$ , and hence  $z(x)$  is lower semicontinuous, but  $\phi(u)$  still jumps from  $-\infty$  to a finite value at  $u = 0$ . Thus boundedness of  $K_1$  is not superfluous for dualizability in Theorem 3. We do not know whether a stochastic program with recourse can exhibit a *finite* duality gap; this is certainly possible for a program of type (1) in which  $f(x)$  is convex and lower semicontinuous but  $K$  is unbounded [3, p. 692].

It can be shown by example that, as the remarks at the end of §2 should suggest, the requirement that  $K_2$  be a polyhedron is not superfluous in Theorem 4. Finally a variant of the example found in [5] will show that finite variance of  $\xi$  is not superfluous in Theorem 4.

REFERENCES

1. R. T. Rockafellar, Duality and stability in extremum problems involving convex functions, *Pacific J. Math.*, 21 (1967), 167-187.
2. -----, Duality in nonlinear programming, in *Mathematics of the Decision Sciences*, Vol. 11 of *Lectures in Applied Mathematics*, American Math. Soc., 1968, 401-422.
3. R. M. Van Slyke and R. J.-B. Wets, A duality for abstract mathematical programs with applications to optimal control theory, *J. Math. Anal. Appl.* 22 (1968), 679-706.
4. D. W. Walkup and R. J.-B. Wets, Stochastic programs with recourse, *SIAM J. Appl. Math.* 15 (1967), 1299-1314.
5. -----, Stochastic programs with recourse: On the continuity of the objective, *SIAM J. Appl. Math.* to appear.
6. -----, A Lipschitzian characterization of convex polyhedra, Boeing document D1-82-0728, Boeing Scientific Research Laboratories, Seattle, Washington 98124.
7. R. J.-B. Wets, Programming under uncertainty: The complete problem, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 4 (1966), 316-339.
8. -----, Programming under uncertainty: The equivalent convex program, *SIAM J. Appl. Math.* 14 (1966), 89-105.